DOCUMENT RESUME

ED 324 349 TM 015 594

AUTHOR	Coley, Richard J., Ed.
TITLE	Testing. ETS Policy Notes, Volume 2, Number 3.
INSTITUTION	Educational Testing Service, Princeton, NJ. Policy Information Center.
PUB DATE	Aug 90
NOTE	13p.
AVAILABLE FROM	ETS Policy Information Center (04-R), Rosedale Road, Princeton, NJ 08541.
PUB TYPE	Collected Works - Serials (022) -- Reports - Evaluative/Feasibility (142)
JOURNAL CIT	ETS Policy Notes; v2 n3 Aug 1990
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Accountability; *Achievement Tests; Constructed Response; *Educational Assessment; Educational Change; Educational Policy; Educational Research; Elementary Secondary Education; Essay Tests; Mathematics Tests; National Programs; Performance; Science Tests; Standardized Tests; *State Programs; Testing Problems; *Testing Programs; Writing Tests
IDENTIFIERS	*Connecticut; Performance Based Evaluation

ABSTRACT

Three articles on current research in testing are presented. The first article, "Testing in the Schools", discusses the role of testing in educational reform. In the 1980s, the overwhelming purpose of state standardized testing has become promoting accountability in areas of: (1) monitoring; (2) gatekeeping; (3) remediation; and (4) funds distribution. Educational policy makers need to find some way to evaluate the tests. Some guideposts for evaluating testing are suggested: making sure that instructional outcomes and learning outcomes guide the testing; determining how a test protects against bias in race, gender, or ethnicity; making sure that appropriate techniques are used; and making testing for accountability less obtrusive. The second article, "Constructed Response Testing: Some Development Efforts", examines two approaches focusing on student-developed solutions to questions that can be economically scored. The first approach involves the use of an answer grid to record answers to mathematical questions, and the second approach involves the use of figural response items for science testing. The third article, "Assessing Performance", describes some of the work conducted at the Educational Testing Service (ETS) and in Connecticut in the area of student performance assessment. The "Learning by Doing" project of the National Assessment of Educational Progress is described. Also discussed are: a writing portfolio study; the Arts PROPEL program in Pittsburgh (Pennsylvania); and Connecticut's Common Core of Learning Assessment Project. Four figures illustrate the discussions. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ETS POLICY NOTES

News from the ETS Policy Information Center

Volume 2, Number 3 Educational Testing Service, Princeton, New Jersey August 1990

ED 324 334

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

RICHARD J. COLEY

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)"

Testing in the Schools

As we begin the 1990s, it is clear that public education and large-scale standardized testing have become interdependent. According to a recent report of the National Commission on Testing and Public Policy, each year elementary and secondary school students take 127 million standardized tests mandated by states and districts, and 20 million school days are devoted to such testing. This averages to about three tests per year per student.

It is not surprising, with such widespread activity, that the desirability and effects of standardized testing are under scrutiny. A number of different concerns about testing have emerged, making it difficult for education officials, policy makers, and the public to achieve focus in discussion and debate. One of the most fundamental concerns is determining the relationship between classroom instruction and standardized tests.

A conventional view is that educators decide what to teach and use a test to see if students have learned what was supposed to be taught. The alternative is to have the test shape instruction. The latter view has been the basis for much of the educational reform during the '70s and early- to mid-'80s; state-mandated tests often evolved from this vision of "tests" as a method for controlling what goes on in the classroom.

Recently, however, arguments have been advanced that the state regulatory approach is too limited, or too centralized, and that nothing less than total "restructuring" is now necessary. As the century closes, testing itself will be tested for its ability to turn around an educational system rated among the lowest performing in the industrialized world.

The Testing Explosion

It is important to understand how standardized testing became the focal point of educational reform. During the 1970s, Americans perceived a decline in educational standards and, consequently, demanded a return to basics. These demands fueled a "minimum competency testing" movement, particularly in reading and mathematics. In their 1982 study, *Measuring the Quality of Education*, Willard Wirtz and Archie Lapointe reported that:

> Thirty-nine states adopted ...minimum competency testing programs. Standardized tests were either developed by state agencies or obtained from commercial publishers. In many cases, specific scores on the examinations were set as marking the lowest levels of competency that could be considered acceptable.

The education system responded to the demand for establishing minimum competencies and succeeded in raising test scores, particularly in lower performance groups. By 1983, however, it was "excellence" that was being demanded, not "minimum competence." In *A Nation At Risk*, the National Commission on Excellence in Education stated, "'Minimum competency' examinations...fall short of what is

This Issue: Testing

One . . . Testing in the Schools
Two . . . Constructed Response Testing: Some Development Efforts
Three . . . Assessing Performance

TM015594

needed, as the "minimum" tends to become the "maximum", thus lowering educational standards for all."

The Commission set a new standard: "Excellence characterizes a school or college that sets high expectations and goals for all learners, then tries in every way possible to help students reach them."

The 1970's wave of standardized testing had been tested and found wanting. Although testing activity intensified with the issuance of *A Nation at Risk* in 1983, mass testing for purposes of school, district and state accountability was not predicted or prescribed by that report. Rather, the report recommended the use of "standardized tests of achievement" for three purposes:

1) to certify the student's credentials
2) to identify the need for remedial intervention
3) to identify the opportunity for advanced or accelerated work

All three purposes relate to the individual student: the first in terms of assessing achievement and the other two as aids for determining the proper course of instruction.

The great majority of states already had statewide testing programs when *A Nation at Risk* was issued. Five years later, in 1989-90, 47 states required that local school districts test public school students at some point or points between grades 1 and 12. This represented an increase of only five states from 1984-85, but during that period many states broadened their testing programs:

- Eleven added new grade levels to be tested, including pre-kindergarten and pre-first grade.
- Six added science and social studies to their testing program, and many more added writing, especially essays, to replace multiple choice exams in language arts.
- Two states moved from testing representative samples of students in a grade to testing all the students.
- Three states switched from allowing local school districts to choose their tests to mandating the use of a state-selected instrument.

While there are a great many distinctions in the scope and purpose of various state testing programs, they can be roughly classified into four categories (see Figure 1). The overwhelming purpose of state testing programs is to promote accountability, and this purpose falls squarely within the regulatory approach to school reform of the 1980s: 38 of the 47 states use tests for monitoring school and/or district performance.

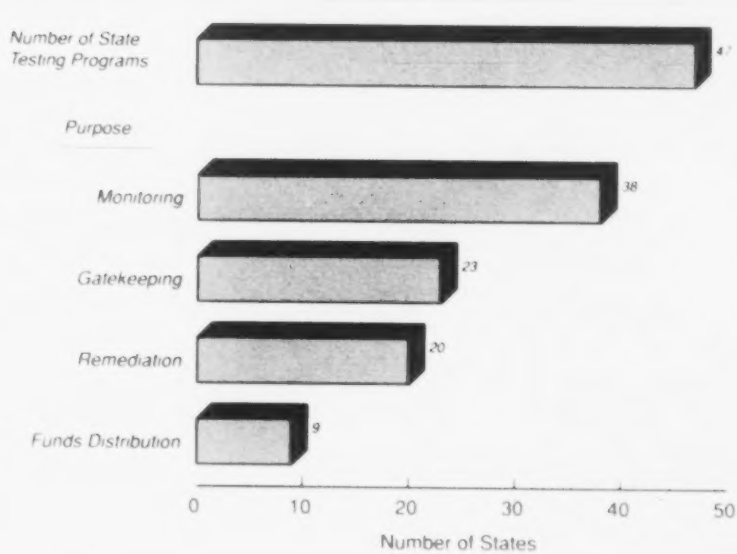
Twenty-three states require tests for grade promotion or high school graduation; another 20 employ testing programs to identify students in need of remediation; and nine states use them in decisions about the distribution of funds.

How Tests Get Used

Of course, from the available data, there is no way to determine how the testing programs are used in the day-to-day practice of education. While some uses are obvious, such as to allocate monetary rewards and penalties to specific schools or to enforce high school graduation requirements, others are more difficult to identify.

When tests are used to monitor education, who is looking at scores, with how much attention, and with what result? When they are used in remedial programs,

Figure 1
State Testing Programs and Purposes, 1990



how many students actually receive different instruction as a result of the tests they have taken, and how is the choice of remedial program guided by the test? Are teachers and schools simply trying to score better on the tests by narrowing instruction, or are they responding by doing a better job of instruction? *While there has been massive testing, there has been minimal study of what actually happens to students, teachers, and schools as a result of this testing.*

In the mid-80s, however, the Center for the Study of Evaluation at University of California at Los Angeles did conduct a national level study dealing with the uses of testing (of all kinds) in the decisions and judgments teachers make. The study reported these findings:

> . . .(the survey results demonstrate that) teachers do use test results of various types in making common instructional decisions. They also reveal quite clearly, however, that teachers place greatest trust in their own observations of students' class performance and in their personal, clinical judgment. Nearly every teacher reporting says that their 'own observations and students' classroom work' are crucial or important sources of information for initially grouping or placing students, in deciding to change students' placement or grouping, and in determining students' report-card grades. The majority also give heavy weight to the results of their own, self-constructed tests in each of these tasks.

Other studies show that teachers do not use the reams of test results they get from the tests that school systems buy. This is the kind of research into actual test use that should be brought to bear on testing decisions and controversies but seldom is.

A Babel of Issues

The discussion about testing lacks focus due to a myriad of issues being addressed by different people with different objectives. The range of concerns includes:

- The ability of present forms of standardized tests to capture thinking and problem solving skills
- The appropriate uses of tests to promote school accountability
- Sole reliance on tests in decision making
- Concern about the adequacy of the "norming" samples frequently used
- Use of too much instructional time for testing
- Sole reliance on multiple-choice testing formats
- The proper fit between what is tested and what is taught
- The appropriateness of stressing the teaching of test-taking skills, "aligning" the curriculum, or actually teaching the test items
- Race, ethnic, or gender discrimination

These issues regarding testing within the K-12 education system differ in some respects from those in admissions testing, where predictive validity is a key matter in judging test quality and appropriateness.

The complications arising from this "Babel" of issues are compounded by the strong emotions that testing often engenders because of personal experiences and the critical ways tests sometimes shape people's lives.

At the same time that the debate is carried out on shifting ground, the standardized testing business is highly technical, with its historical grounding in psychology and statistics. Few engaged in the debate understand such esoteric terminology as item response theory, domains, questions of dimensionality, and norm and criterion referencing, to name a few. And common terms such as validity and reliability have taken on very technical meaning in testing. Now, the terms "authentic assessment" and "performance testing" are entering the discussion.

A Few Guideposts

It is amidst this bewilderment and confusion that educators, policy makers, and the public must form opinions and make decisions that will affect the future of our education system. These are decisions that should not be delegated to technicians. It is becoming more and more apparent that to make critical choices about elementary and secondary education, we need to find or develop methods to "test" the testing programs. Unfortunately, tests are used for so many different purposes in the education system that no single set of rules will suffice. There is now emerging, however, some consensus within what might be called "the educational testing reform movement" as to the general directions testing needs to go if educational objectives are to be achieved. While the "guideposts" suggested below are this author's formulation, they

are believed to parallel what this reform movement is generally saying. (An additional source of guidance for educators, policy makers, and others is *The Code of Fair Testing Practices in Education*, issued by the American Educational Research Association, the National Council on Measurement in Education, and the American Psychological Association and publicly endorsed by some leading publishers of tests, including ETS.)

Make Sure that Instructional Objectives and Learning Outcomes Guide the Testing

> Standardized achievement tests have no intrinsic value in education other than as measures of whether instruction has had its intended effects. The educational enterprise determines what students should be taught and what they should know. Make the tests fit the instructional goals and strategies, not vice versa. Every test embodies an implicit theory of how learning does or should occur. Make sure that theory is known.

Know How a Test Protects Against Bias in Terms of Race, Ethnicity, and Gender

> Test scores should reflect what students learn. Test constructors must guard against scores being affected by characteristics extraneous to instructional objectives and outcomes. How does the test development process guard against bias? After being satisfied as to cultural fairness, use score differences among groups of students to determine where to focus more instructional attention.

Make Sure that the Techniques of Testing Are Appropriate for Measuring Desired Knowledge and Skills

> This issue has come to the fore recently with criticism of sole reliance on multiple-choice formats in standardized testing. Terms such as "authentic," "constructed response," and "performance" testing are appearing more frequently. The debate about the effects of multiple-choice test formats still requires substantial research about the skills it is best suited to measure. We need to know more about how much difference there is, in terms of what is measured, between choosing among answers and constructing them. Some research on this matter is available, much of it conducted by researchers at Educational Testing Service. More knowledge is needed about the best applications for a variety of testing formats, as is more effort to establish the measurement characteristics of performance tests. Serious desire for open-ended questions and performance measures must be matched by serious attention to the time and resource implications that are involved.

Make Testing for Accountability Less Intrusive

> Every *ten* years the nation conducts a census. Its regular information comes from carefully constructed national samples of households, carried out by carefully trained interviewers, using instruments carefully developed and tested over many years. This is also the means by which the widely respected National Assessment of Educational Progress gathers its data about achievement. This type of sampling system could be used for accountability testing. It would intrude less on valuable class time and would not interfere as much with legitimate instructional objectives determined by schools and districts. The use of sampling systems for evaluation also would permit a lot more flexibility to introduce formats other than multiple-choice. Of course, while this can be done to evaluate the performance of the system, only individual testing can be used to satisfy a high school graduation requirement or to inform an instructional decision about an individual student.

Last February President Bush and the nation's governors announced National Goals for Education, to be achieved by the year 2000. They are very ambitious goals. Testing and assessment are interwoven with educational practice and management, and complementary goals will need to be established. To be useful, such goals will have to emerge from a very thoughtful process, one that is informed by relevant research and analysis of existing programs, not by the slogans arising from heated debates about testing.

Constructed Response Testing: Some Development Efforts

There is a growing view that tests in elementary and secondary schools should not rely so heavily on multiple-choice formats, but there is also the reality that multiple choice testing has an incomparable advantage from a cost standpoint: the answer sheets can be machine scored. In complete open-ended question formats, each paper has to be graded by a human being and under carefully controlled conditions that assure that uniform scoring standards are applied. Two separate development efforts at Educational Testing Service have explored approaches that focus on student-developed solutions to questions that can be economically scored. The first involves the use of a "grid" for recording answers to mathematics questions. The second uses "figural response items" in science, where the examinee is called upon to complete a partially completed figure, or otherwise indicate something on a drawing or graph. These examples are among a considerable number of new approaches to testing formats being developed at ETS.

An Answer Grid

ETS has developed and tested two prototypes in which multiple-choice mathematics questions were converted, so that the answer could be recorded in a grid that can still be machine scored. In one prototype, the multiple-choice and the grid versions were each given to equivalent samples of over 900 high school juniors and seniors (see Figure 2 for an example from each test).

Figure 2

The Question:

Section I of a certain theater contains 12 rows of 15 seats each. Section II contains 10 rows, but has the same total number of seats as Section I. If each row in Section II contains the same number of seats, how many seats are in each row?

Test 1, Multiple Choice Version

(A) 16
(B) 17
(C) 18*
(D) 19
(E) 20

Test 2, Grid Version

In comparing the results, ETS found that the grid questions "worked well" from a test construction point of view, and that the results were not differentially related to gender. However, the "gridding in" test was harder: the average percent correct for the multiple-choice version was 54.5 percent, compared to 47.4 percent for the grid version.

In one test, students were given both multiple-choice and "grid-in" questions and asked to make some comparisons. Seventy-six percent thought the multiple-choice questions were easier to answer. Fifty-seven percent thought the grid-in questions were "a better measure" of their ability in mathematics. Twenty-two percent thought multiple-choice questions were a better measure, and 21 percent saw no difference.

In a second stage of the project, a more elaborate prototype was developed, making it possible for students to grid fractions, decimals, and whole numbers.

While the use of grid-in type items seems to be operationally feasible, based on research to date, James Braswell of ETS states that "additional study is needed to refine the directions and the format and to determine the appropriate timing and difficulty level for a group of items in this format." It is also necessary to determine what age groups are appropriate for use of the grid format.

Figural Responses

An alternative to multiple-choice questions in science testing is to use a drawing which is not complete, or on which an examinee can mark a location (Figure 3). For example, examinees may be asked to indicate a direction by drawing in arrows or to show the location of an anatomical flaw in a diagram of a heart. While this approach to free response is one that does not depend on multiple-choice questions, the figures are designed so that the answers can still be machine scored.

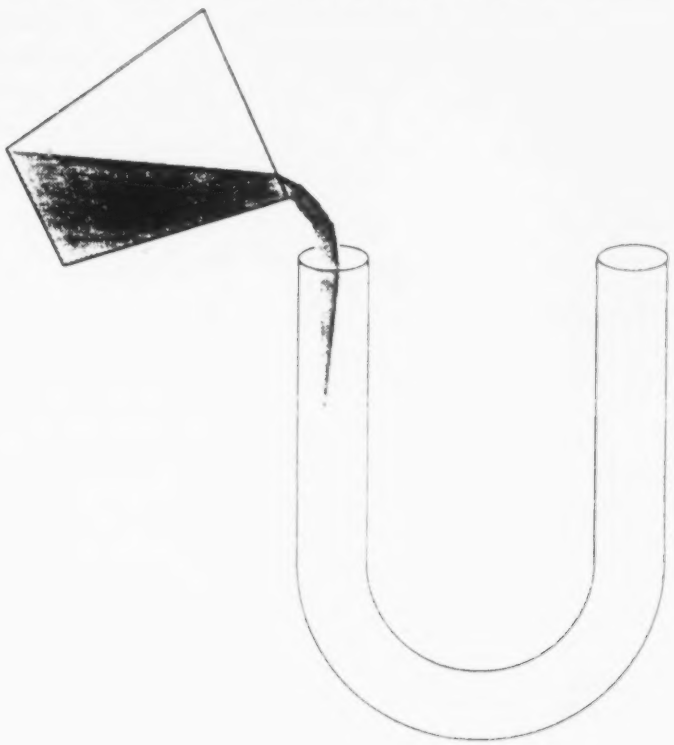
A study using such "figural response" questions was part of the field test for the development of the 1990 National Assessment of Educational Progress (NAEP). The results of the field test disclosed that the figural response items were, in general, more difficult than their multiple-choice counterparts. The figural response questions were used with samples of students in grades 4, 8, and 12, with the samples designed to represent a broad range of student characteristics, such as racial/ethnic group membership and socioeconomic status. Sixteen of the figural response items tested were used by NAEP.

• • • •

Both these descriptions are taken from presentations made at the April, 1990, meeting of the American Educational Research Association. The first, on grids, was presented by James Braswell at Educational Testing Service; this paper is being revised and will be in final preparation, available in the Fall of 1990. The second, on figural response items, was by Michael E. Martinez, also of Educational Testing Service. His paper is titled "A Comparison of Multiple-Choice and Constructed Figural Response Items," April, 1990. ■

Figure 3

On the diagram below, draw where you think the water level would be after all the water in the beaker is poured into the U-shaped tube.



Assessing Performance

The terms "problem solving," "critical thinking," and "higher order skills" are becoming pervasive in discussions about education reform. The common belief is that such skills are imparted in more active learning environments, or through "hands-on" approaches. Discussions of how students learn in this mode lead to questions about how to assess what they learn. The discussion moves in two directions: from instruction to the question of how to assess results and from the design of a performance assessment that will drive a "hands-on" approach to the instruction necessary to prepare for it. We briefly report on some of the pioneering work in assessing performance at Educational Testing Service (ETS) and in the state of Connecticut — two of the many places where performance-based measures are being tried and used.

The NAEP "Learning by Doing" Project

The National Assessment of Educational Progress (NAEP), administered by ETS, has developed and pilot-tested a variety of hands-on science and mathematics tasks to be used as prototypes for future assessments.* The tasks required students to think independently about a variety of relationships. Here are a few examples.

At the first level, students are asked to classify and sort by identifying common characteristics of plants and animals.

- At grades 7 and 11, students were asked to sort a collection of small-animal vertebrae into three groups and explain how the bones in each grouping are alike.

At the next level, students are given materials and asked to observe, infer, and formulate hypotheses.

- At grades 3 and 7, students are asked to describe what happens when a drop of water is placed on different types of building materials and then apply what they have learned by hypothesizing what the water will do when placed on an unknown material.

At the most complex level, students are asked to design and conduct complete experiments.

- At grade 11, students are asked to design a reliable experiment to determine the effects of exercise on heart rate. Students need to identify the variables to be manipulated, specify what needs to be measured, and describe how the measurements should be made. (This exercise was included as a prototype to assess students when actual experimentation in a classroom or assessment setting is difficult.)

The results of the pilot test were encouraging. Although managing equipment and training administrators required ingenuity and painstaking effort, the project showed that conducting hands-on assessment is feasible and extremely worthwhile. Professional educators were enthusiastic; students were engaged by the tasks; and school staff encouraged further use of these kinds of tasks in both instruction and assessment. However, NAEP has not yet been funded to actually carry out such a "hands-on" assessment with a national sample of students.**

Writing Portfolios

NAEP staff at ETS have designed an experimental writing portfolio study to be conducted this year that will permit an evaluation of writing that students produce for their school assignments, rather than within a testing situation. The project specifies that teachers submit students' writing that has been produced in school for an assignment. In addition to providing more extensive writing samples for assessment than is possible in a testing situation, the hope is that the portfolio materials will provide some information about the kinds of writing tasks being assigned in the nation's classrooms. The portfolio approach offers an opportunity to use the best of current knowledge about writing theory and instruction in the design and implementation of more appropriate forms of writing assessment. It also implements on a national scale some of the innovative writing portfolio efforts

* NAEP is administered under a contract with the U.S. Department of Education's National Center for Education Statistics (NCES). The Learning by Doing Project was funded by the National Science Foundation through a grant to NCES.

** In response to the positive results of the pilot study, *Learning by Doing* was published to describe the tasks field tested by NAEP. It is available from NAEP, P.O. Box 6710, Princeton, NJ 08541-6710 for $5 plus $1.50 shipping and handling. *Learning by Doing* was adapted from *A Pilot Study of Higher-Order Thinking Skills Assessment Techniques in Science and Mathematics: Final Report*. This two-volume, 537-page report, which describes NAEP's project in detail and presents all 30 tasks included in the pilot study is available for $35 plus $1.50 shipping and handling from the address above.

currently being undertaken in states such as Rhode Island and Vermont. This approach was incorporated into the 1990 NAEP Assessment of Writing and will be continued in the 1992 assessment.

During the last few years, ETS personnel have pioneered a variety of portfolio applications. Particularly noteworthy are the contributions of Mary Fowles and Roberta Camp. In Rhode Island, for example, ETS staff worked with the state education department, the Rhode Island Consortium on Writing, and Rhode Island teachers to develop a portfolio-based program to test the validity of the state's earlier assessment of third grade writing.

Arts PROPEL

Creating curricula and assessments that provide a much richer depiction of how children learn music, visual arts, and creative writing is the goal of Arts PROPEL, a project involving the Pittsburgh Public Schools, ETS, and Harvard University's Project Zero. The project is funded by the Rockefeller Foundation. PROPEL is designing assessments that are woven into daily classroom instruction. As students produce sketchbooks and journals, compile portfolios, and complete carefully sequenced classroom activities, they leave behind a series of "footprints" for teachers about how they are growing and thinking as artists. The project's hope is that the more rapid, qualitative feedback provided by these assessments will prove more meaningful to students and teachers than current tests. In addition, since the exercises double as instructional tools, they are helping to modify the curriculum.

Reflective interviews are an example of one technique used to allow students to judge themselves. As part of the process of reflecting on the body of their work, students can become aware of the particular signature they give to prints, performances, or poems. This becomes evident in the case of Connie, a high school junior who, in the course of her writing, turned out a series of short poems. When asked by her teacher to reflect on her writing the same way she might think about poems by Yeats or Dickinson, Connie noticed — for the first time — that she had a style, a characteristic signature, as a writer. She was able to see how consistently she dealt with the hard facts and small ironies of everyday life by making common objects, like mops, speak (Figure 4).***

Figure 4

A Sample of Connie's Poetry

Mop

Woman tall and thin
With long tangled gray hair
Must turn her life upside down
To do her duty
Hold her breath while washing
her hair
Wringing out the dirty water
Then she goes to her duty
Again.

Performance Assessment in Connecticut

The state of Connecticut is active on several fronts in the development of performance assessment.**** The Connecticut Multi-State Performance Assessment Coalition Team (COMPACT) Project, sponsored by the Connecticut State Department of Education and the National Science Foundation, is a collaboration of the State Departments of Education from Connecticut, Michigan, Minnesota, New York, Texas, Vermont, and Wisconsin; the Coalition of Essential Schools (CES); The Urban District's Leadership Coalition of the American Federation of Teachers; and Project Re:Learning.

Connecticut's Common Core of Learning Assessment Project assesses high school math and science students working together in groups to solve problems and design and conduct experiments. This fits with a view of students as knowledge workers, whose job it is to construct meaning from what they know and the new information they encounter. The teacher's role is to be the manager of these knowledge workers.

The Project will use a Core of Learning exam, expected to be in place in 1991. It is designed to force students to think before answering. Here's an abbreviated version of one test. The problem is: How can you really tell which food market will save you the most money? Your assignment: Design and carry out a study to answer

*** Wolf, Dennie Palmer. "Opening Up Assessment." *Educational Leadership*, December 1987/January 1988.

**** For more information about these new projects in Connecticut, contact Joan Boykoff Baron, Connecticut Common Core of Learning Coordinator, Connecticut State Department of Education, Box 2219, Hartford, CT 06145.

the problem. The Project takes several steps.

1. Write a report that outlines how you would solve the problem. What markets will you compare? What items? How and why did you make your choices? What records will you keep? How will you analyze the data? Keep a log reporting the progress of your project.
2. Form a research group with 3-4 people. You will meet twice in class to compare your plans and to develop a final, written research approach. Hand it in for comments and grading.
3. Carry out the study, with each group member doing a portion of the work. Hand in a final report as a group. The report should restate the problem that was solved, explain how the data were collected and analyzed, and include graphics that will illustrate your conclusions.*****

* * *

These few examples provide a sampling of new work being undertaken at ETS and at other places across the country in developing alternative methods of performance assessment. They are illustrative of new attempts to go beyond traditional assessment methods.

"INTELLIGENT" ASSESSMENT

"Intelligent Assessment is conceived of as an integration of three research lines, each dealing with cognitive performance from a different perspective: constructed response testing, artificial intelligence, and model-based assessment. This integration is envisioned as producing assessment methods consisting of tasks closer to the complex problems typically encountered in academic and work settings . . . It is important to stress that the emphasis is on assessment that facilitates instruction . . ."

Randy Bennett, Educational Testing Service, 1990

TESTING TO "FACILITATE SUCCESS"

"New developments in measurement, especially in concert with new developments in cognitive and computer science, afford both new reasons and new possibilities for developing direct measures of student performance. Performance processes would be assessed directly by means of work samples or simulations of real-world generic tasks, rather than in terms of total scores summarizing the piecemeal information provided by a set of discrete test items."

Samuel Messick, Educational Testing Service, 1988

***** *Newsweek*, January 8, 1990, p. 58.

ETS Policy Information Center Publications

ETS Policy Notes Newsletters

Vol. 1, No. 1, July 1988
"Black College Faculty: A Dwindling Resource"
"Introducing the ETS Policy Information Center"
"Who's Going to Graduate and Professional Schools?"
"What's Wrong With This Picture?"
"State Profiles of Educational Standards Updated"
"Report Highlights College Minority Retention Programs"
"New Studies Monitor Talent Flow Into Technical Fields"

Vol. 1, No. 2, March 1989 — From High School to College
"Edging Forward: What the SAT Shows About College-Bound Seniors in the 1980s"
"Starting on the Right Track"
"High-Achieving Hispanic Students"

Vol. 1, No. 3, June 1989 — Science
"A Precious Few: Interest of the College-Bound in the Quantitative Sciences"
"A Straggler's View: The U.S. in the World of Science Education"
"Staying Power: Students Who Persist"

Vol. 2, No. 1, October 1989 — The Gender Gap
"The Gender Gap in Education: How Early and How Large?"
"Scholastic Ability"
"Sex Differences in Test Performance: A Synthesis of Research"

Vol. 2, No. 2, March 1990 — Public School Choice
"Choice in Montclair, New Jersey"
"What the Research Says"
"Update on State Activity"

Vol. 2, No. 3, August 1990 — Testing
"Testing in the Schools"
"Constructed Response Testing: Some Development Efforts"
"Assessing Performance"

(Available (while supplies last) from ETS Policy Information Center (04-R), Rosedale Road, Princeton, NJ 08541.)

Reports

Skills Employers Need: Time to Measure Them? A Policy Information Proposal. June 1990

(Available for $3.50, prepaid, from ETS Policy Information Center (04-R), Rosedale Road, Princeton, NJ 08541.)

This brief paper summarizes the skills that employers want from job candidates and proposes the development of an Employment Readiness Profile. This profile would provide a barometer of progress in producing a quality labor force.

From School to Work. A Policy Information Report. 1990

(Available for $3.50, prepaid, from ETS Publications Order Service, P.O. Box 6736, Princeton, NJ 08541-6736. Order No. 204840)

The U.S. is among the worst in the industrial world in helping students who don't go on to college make the transition from school to work. This report discusses student work during high school, differences between skills acquired in the classroom and those needed at the workplace, the information processing skills of high school graduates, new efforts to integrate academic and vocational education, and the weakness of linkages between the school and the workplace.

Choice in Montclair, New Jersey. A Policy Information Paper. Beatriz C. Clewell and Myra F. Joy. January 1990.

(Available for $5.00, prepaid, from ETS Policy Information Center (04-R), Rosedale Road, Princeton, NJ 08541.)

Montclair, New Jersey, is an urban school district that has achieved success in desegregating its schools through a voluntary magnet school plan based on choice. To study the effectiveness of Montclair's plan in providing racial balance across schools and educational quality and diversity in programs through the use of choice, the authors conducted a case study of the district in 1987 and a follow-up in the summer of 1989. The paper reviews a variety of public school choice programs and describes and evaluates the Montclair model. The paper outlines the factors contributing to the district's success and offers some recommendations concerning the development and implementation of similar public school choice plans.

What Americans Study. A Policy Information Report. 1989

(Available for $3.50, prepaid, from ETS Publications Order Service, P.O. Box 6736, Princeton, NJ 08541-6736. Order No. 204834)

Increasing course requirements in key academic subjects has been a central theme of educational reform in the decade of the 1980s. This report provides information on what is being studied and on how this has changed over time for high school graduates and college-bound seniors. It also describes course-taking patterns for eleventh-, eighth-, and fourth-grade students.

State Education Indicators: Measured Strides, Missing Steps. Stephen S. Kaagan and Richard J. Coley. 1989.

(Available for $3.75, prepaid, from ETS Publications Order Service, P.O. Box 6736, Princeton, NJ 08541-6736. Order No. 239012)

The monograph describes the central features of indicator systems and the issues that must be addressed with regard to their purposes, applications, and effects at the state and local levels. It also provides case studies of state education indicator systems in California, Connecticut, New York, and South Carolina.

Earning and Learning. Paul E. Barton. March 1989.

(Available for $3.50, prepaid, from the National Assessment of Educational Progress, Educational Testing Service, Rosedale Road, Princeton, NJ 08541-0001. Order No. 17-WL-01)

This report explores the relationship between work and student achievement, using information from the 1986 National Assessment of Educational Progress (NAEP). It relates hours worked per week to student achievement on the NAEP proficiency scale for each subject area assessed. It describes who works and who does not, examines the adjustments working students make in other activities, charts the growth of the student work force, and summarizes the results of major research projects that have addressed the effects of student work on school performance.

Information for National Performance Goals for Education: A Workbook. November 1989

(Available for $3.50, prepaid, from ETS Policy Information Center (04-R), Rosedale Road, Princeton, NJ 08541.)

This "workbook" was prepared to assist those charged with setting national education performance goals as a result of the Education Summit held by President Bush and the nation's governors in Charlottesville, Virginia. It assembles information about current and past educational performance to inform decisions about outcome goals for the future.

A Conference on Construction vs. Choice in Cognitive Measurement

Sponsored by

Educational Testing Service

November 30 & December 1, 1990

Topics will include:

Studying Differences Between Multiple Choice and Free Response

•

Non-Test-Based Approaches to Cognitive Assessment

•

Test Development and Scoring Issues

•

The Politics of Free Response vs. Multiple Choice

For information contact:
Ms. Terri Sterling • ETS • Mailstop 20-T •
Princeton, NJ 08541 • (609) 734-1550

ETS Policy Notes is published by the ETS Policy Information Center, Educational Testing Service, Princeton, NJ 08541-0001; (609) 734-5694.

Director, ETS Policy Information Center: Paul E. Barton

Editor: Richard J. Coley

Copyright © 1990 by Educational Testing Service. All rights reserved. Educational Testing Service is an Affirmative Action/Equal Opportunity Employer.

Educational Testing Service, ETS, and ® are registered trademarks of Educational Testing Service.

04702 - 02536 • 4B0M6 • 204839 • Printed in U.S.A.

FIRST CLASS
EDUCATIONAL
TESTING SERVICE
U.S. POSTAGE
PAID